Targeted genomic rearrangements using CRISPR/Cas technology

Published in final edited form as:
Nat Commun. ; 5: 3728. doi:10.1038/ncomms4728.

Targeted genomic rearrangements using CRISPR/Cas technology

Peter S. Choi[1,2] and **Matthew Meyerson**[1,2,3,4]

[1]Department of Medical Oncology, Dana-Farber Cancer Institute, Boston, Massachusetts, USA

[2]Broad Institute of MIT and Harvard, Cambridge, Massachusetts, USA

[3]Department of Pathology, Harvard Medical School, Boston, Massachusetts, USA

Abstract

Genomic rearrangements are frequently observed in cancer cells but have been difficult to generate in a highly specific manner for functional analysis. Here we report the application of CRISPR/Cas technology to successfully generate several types of chromosomal rearrangements implicated as driver events in lung cancer, including the *CD74-ROS1* translocation event and the *EML4-ALK* and *KIF5B-RET* inversion events. Our results demonstrate that Cas9-induced DNA breaks promote efficient rearrangement between pairs of targeted loci, providing a highly tractable approach for the study of genomic rearrangements.

Introduction

Cancer cells are characterized by abnormalities in chromosome number and structure, such as genomic rearrangements that can result in the inappropriate expression or activation of oncogenes[1]. While many novel cancer-associated rearrangements continue to be identified, it remains difficult to accurately model these events for functional analyses. The ability to efficiently generate specific rearrangements would greatly improve our understanding of how structural variations in the genome arise and how each of these events contributes to disease pathogenesis.

The CRISPR/Cas system has recently been adapted to provide site-specific DNA recognition and cleavage through a customizable RNA guide[2, 3, 4, 5]. Cas9 from *Streptococcus pyogenes* recognizes a 20 nucleotide (nt) target sequence immediately upstream of the requisite protospacer-adjacent motif (PAM) sequence NGG[6]. Co-expression of the Cas9 enzyme and a chimeric single-guide RNA (sgRNA) results in Cas9-induced double-strand breaks (DSBs) at the targeted genomic sequence[6] (Fig. 1a). We hypothesized that the high efficiency of DNA cleavage mediated by Cas9 would facilitate the formation of rearrangements in a targeted manner.

[4]Corresponding author matthew_meyerson@dfci.harvard.edu.

Author contributions P.S.C. and M.M. designed the research. P.S.C. performed the experiments and analyzed the data. P.S.C. and M.M. wrote the manuscript.

Competing financial interests The authors declare no competing financial interests.

Here we specifically investigate whether pairs of DSBs induced by RNA-guided Cas9 would be sufficient to generate chromosomal translocations and inversions (Fig. 1a). We model several genomic rearrangements, known as driver events, in lung adenocarcinoma which represent a variety of rearrangement types. We find that DNA cleavage by Cas9 at two genomic loci results in detectable levels of rearrangement between the targeted regions. DNA rearrangement also results in expression of the expected fusion transcripts and protein products, demonstrating that CRISPR/Cas technology is a highly practical tool for the study of genomic rearrangements.

Results

In lung adenocarcinoma, *ROS1* is involved in translocations that result in in-frame fusions with *CD74* or *SLC34A2*[7, 8, 9]. We first attempted to generate the *CD74-ROS1* rearrangement, which arises through a translocation between *CD74* on chromosome 5 and *ROS1* on chromosome 6 (Fig. 1b). We designed sgRNAs targeting intron 6 of *CD74* and intron 33 of *ROS1* (Fig. 1b), which were then co-expressed with Cas9 in HEK 293T cells. Cleavage of each targeted region was highly efficient, as assessed by the formation of indels using the Surveyor assay (Supplementary Fig. 1a). Using primers spanning the expected breakpoint junction, we detected translocations occurring in cells expressing both *CD74* and *ROS1* sgRNAs, but not in cells expressing only a single targeting sgRNA (Fig. 1c,d). Sequencing of breakpoints confirmed formation of the *CD74-ROS1* translocation event and we observed junction types resulting from both precise joining of predicted cleavage sites as well as those containing short deletions that likely result from nucleolytic processing of DNA ends during DSB repair (Supplementary Fig. 2a,b). In addition, we detected expression of the predicted *CD74-ROS1* fusion transcript from cDNA samples using primers spanning the junction between *CD74* exon 6 and *ROS1* exon 33 (Fig. 1e). We were also able to generate the same *CD74-ROS1* translocation in non-transformed immortalized lung epithelial cells (AALE)[10], which represent a more relevant cellular context for studying the *CD74-ROS1* rearrangement event (Supplementary Fig. 3). Collectively, these results demonstrate that Cas9-induced DSBs are sufficient to promote translocations between targeted chromosomes in multiple cell types.

Next, we sought to determine whether Cas9-mediated DNA cleavage could promote the formation of intrachromosomal inversions. We chose to model the *EML4-ALK* and *KIF5B-RET* rearrangements observed in lung cancer, which represent two different types of inversion events. *EML4-ALK* is the result of a paracentric inversion of chromosome 2p[11], while *KIF5BRET* is the result of a pericentric inversion across the two arms of chromosome 10[9, 12, 13, 14] (Fig. 2a,f). We first selected target sites that were in the immediate vicinity of intronic breakpoints identified in patient samples containing the *EML4-ALK*[11] and *KIF5B-RET*[13] rearrangements. Introduction of constructs expressing Cas9 and each targeting sgRNA into HEK 293T cells resulted in efficient site-specific cleavage, as detected by indel formation (Supplementary Fig. 1b,c). For both *EML4-ALK* and *KIF5B-RET*, the expected inversions were only detected in cells expressing Cas9 along with the appropriate pair of sgRNAs, as compared to controls cells expressing only one sgRNA (Fig. 2 b,g). *EML4-ALK* and *KIF5B-RET* fusion transcripts were also detected only when both sites were targeted (Fig. 2 c,h), and matched the predicted sequence (Fig. 2 d,i; Supplementary Figure 4). In

addition, among single-cell clones that were derived from cells transfected with the combination of *EML4* and *ALK* sgRNAs, we identified several positive clones that expressed the EML4-ALK fusion protein (Fig. 2e).

To estimate the efficiency of the induced *EML4-ALK* and *KIF5B-RET* rearrangements, we developed a flow cytometric assay to quantify ALK or RET protein expression at the single-cell level. Untreated HEK 293T cells do not express appreciable amounts of either ALK or RET, allowing the levels of these proteins to serve as surrogate markers of successful rearrangement. For the *EML4-ALK* rearrangement, we observed increased ALK expression in approximately 8% of cells treated with both *EML4* and *ALK* sgRNAs, compared to 1% or less of cells treated with single sgRNAs or no sgRNAs (Fig. 3a), while for the *KIF5B-RET* rearrangement, we observed increased RET expression in approximately 1.6% of cells treated with both *KIF5B* and *RET* sgRNAs, compared to less than 0.5% of cells treated with single sgRNAs or no sgRNAs (Fig. 3b). Taken together, our results demonstrate that intrachromosomal inversions involving a single arm or occurring across both arms can be readily generated using the CRISPR/Cas system.

Discussion

We have found that targeted DNA DSBs induced by Cas9 can promote genomic rearrangements of several types, including interchromosomal translocations and intrachromosomal inversions. Previously, it has been reported that chromosomal rearrangements can also be generated using meganucleases, zinc-finger nucleases (ZFNs) and transcription activator-like effector nucleases (TALENs)[15, 16, 17]. While all of these tools have revolutionized our ability to manipulate the genome, the CRISPR/Cas9 system offers several unique advantages. The ease with which sgRNAs can be designed and cloned makes the CRISPR system well-suited for high-throughput screening experiments. Also, the ability to multiplex Cas9 targeting opens up the possibility of studying more complex combinations of rearrangements observed in cancer genomes, such as those resulting from chromothripsis[18] or chromoplexy[19]. In addition, while off-target recognition and cleavage by Cas9 may result in non-specific chromosomal rearrangements, these unintended effects could be minimized by generating DSBs using a ‘double-nicking’ strategy[20, 21] or targeting with shorter 17nt sgRNAs which exhibit greater specificity[22].

The rearrangements and inversions induced in this study are most likely occurring through the non-homologous end-joining (NHEJ) pathway of DSB repair, as they involve the joining of mismatched ends. We did not detect regions of significant homology between the targeted DNA regions, suggesting that alternative forms of repair, such as non-allelic homologous recombination (NAHR) are unlikely to be responsible for generating the induced rearrangements. From our analysis of the rearranged genomic breakpoint junctions, we observed a high frequency of exact fusion events, which may be explained by the mechanism of Cas9-mediated DNA cleavage. It has been described previously that Cas9 generates blunt ends precisely three base-pairs upstream of the PAM sequence[6]. Since the resulting blunt DNA ends would not require further end-processing prior to ligation in the NHEJ pathway of repair, rearrangements between Cas9-generated ends may frequently occur as exact fusion events. Similarly, NHEJ-mediated repair of blunt-ended signal joints

generated during V(D)J recombination also results in precise error-free fusions[23]. However, the nature of an induced fusion site is also likely to depend on a variety of other factors, such as features specific to the genomic regions that are targeted, the unique properties of the cell lines that are used, and the levels of Cas9/sgRNAs that are expressed.

We propose that Cas9-mediated targeting of endogenous loci provides a highly useful approach for the study of novel chromosomal rearrangements and their functional significance in human disease. Accurate modeling of rearrangements will provide opportunities to not only study the phenotypic consequences of specific events, but to also investigate the mechanisms that contribute to structural changes to the genome.

Methods

CRISPR sgRNA target selection and cloning

Suitable sgRNA target sites were manually identified based on the following criteria: an initiating G on the 5′ end, the presence of an NGG motif on the 3′ end, proximity to breakpoints previously reported in patient samples (when such information was available), and sequence uniqueness by BLAST search. Oligos encoding the targeting sequence were then annealed and ligated into the BbsI sites of pX330[3] (Addgene #42230).

Cell transfection and DNA extraction

HEK 293T cells were plated the day prior to transfection at 10^6 cells per well of a 6-well plate. On the day of transfection, the media was replaced to antibiotic-free DMEM/10% FBS. Cells were treated with 25μM of chloroquine and transfected with a total of 2μg of DNA by the calcium phosphate method (CalPhos Kit, Clontech #631312). Transfection efficiencies were estimated by transfection of a GFP reporter plasmid in each experiment and averaged 85-95%.

For transfection of AALE cells, 5×10^5 cells were nucleofected with a total of 2μg of DNA of DNA using the Basic Epithelial Cell Nucleofector Kit (Lonza VPI-1005) and the T13 setting. Media was replaced the following day and cells were harvested approximately 72hrs post-transfection. DNA was extracted from cells using Quickextract DNA extraction solution (Epicentre #QE09050) and diluted to 10ng μl^{-1} for subsequent PCR analyses.

RNA extraction and RT-PCR

Total RNA was purified from cells using the NucleoSpin RNA kit (Machery-Nagel #740955). First-strand cDNA synthesis was subsequently carried out on 1μg of total RNA with the iScript cDNA Synthesis Kit (Bio-Rad #170-8891).

PCR-based detection of genomic breakpoints and cDNA fusions

PCR amplification was carried out using Q5 high-fidelity DNA polymerase (NEB #M0491). PCR products were electrophoresed on 2% TBE gels and purified using the Nucleospin Gel and PCR Clean-up kit (Machery-Nagel #740609). Purified PCR products were directly sequenced using the forward PCR primer. *CD74-ROS1* and *ROS1-CD74* genomic breakpoint PCR products were cloned into pJET1.2 using the CloneJET PCR Cloning Kit

(Thermo Scientific #K1231) and individual clones were sequenced. PCR primers are listed in Supplementary Table 1.

Isolation of EML4-ALK clones and western blotting

HEK 293T cells transfected with both *EML4* and *ALK* sgRNAs were single cell cloned by limiting dilution in 96-well plates. Individual clones were screened by PCR for the *EML4-ALK* genomic rearrangement. PCR-positive clones were subsequently screened by western blotting for ALK expression using standard procedures. Cells were lysed in RIPA buffer and lysates were separated on 4-12% Bis-Tris NuPage gels (Invitrogen). After transferring to PVDF membranes, blots were probed with antibodies against ALK (Cell Signaling #3633, 1:1000) or ACTB (Cell Signaling #8457, 1:1000).

Surveyor nuclease assay

Genomic regions surrounding sgRNA-targeted sites were amplified by PCR (see Supplementary Table 1 for primer sequences). PCR products were gel-extracted or purified directly using the Nucleospin Gel and PCR Clean-Up Kit (Machery-Nagel, #740609). Approximately 200-400 ng of each PCR product was then denatured and re-annealed in 20μl total volume containing Pfu Turbo reaction buffer, as specified in the Surveyor Mutation Detection Kit protocol (Transgenomics #706020). Next, 2μl of 0.15M $MgCl_2$, 1μl of Surveyor enhancer solution, and 1μl of Surveyor nuclease was added and the mixture was incubated at 42C for 60min. Reactions were stopped with addition of 2μl of Stop solution and separated on 2% TBE agarose gels.

Intracellular staining for ALK and RET

Cells were washed in PBS and fixed with 2% paraformaldehyde in PBS for 10 min. at 37°C. After washing in PBS, cells were permeabilized in cold 90% methanol for 30 min. on ice. Cells were then washed in 0.5% BSA in PBS and stained with anti-ALK (Cell Signaling Technology, #3633P, 1:400) or anti-RET (Cell Signaling Technology, #3223S, 1:50) for 1 hour at room temperature, followed by a wash in 0.5% BSA in PBS, and staining with anti-rabbit Alexa Fluor 488 secondary antibody (Life Technologies, #A11070, 1:500) for 30 min at room temperature. After a final wash, cells were resuspended in PBS and analyzed on a BD LSR Fortessa flow cytometer.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank Hideo Watanabe and other members of the Meyerson laboratory for helpful suggestions, and Feng Zhang for advice and reagents. P.S.C. is supported by a fellowship from the International Association for the Study of Lung Cancer. This research was supported by NCI R01CA109038 and DOD grant W81XWH-12-1-0269 (MM).

References

1. Mitelman F, Johansson B, Mertens F. The impact of translocations and gene fusions on cancer causation. Nature reviews Cancer. 2007; 7:233–245.

2. Cho SW, Kim S, Kim JM, Kim JS. Targeted genome engineering in human cells with the Cas9 RNA-guided endonuclease. Nature biotechnology. 2013; 31:230–232.
3. Cong L, et al. Multiplex genome engineering using CRISPR/Cas systems. Science. 2013; 339:819–823. [PubMed: 23287718]
4. Jinek M, East A, Cheng A, Lin S, Ma E, Doudna J. RNA-programmed genome editing in human cells. eLife. 2013; 2:e00471. [PubMed: 23386978]
5. Mali P, et al. RNA-guided human genome engineering via Cas9. Science. 2013; 339:823–826. [PubMed: 23287722]
6. Jinek M, Chylinski K, Fonfara I, Hauer M, Doudna JA, Charpentier E. A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity. Science. 2012; 337:816–821. [PubMed: 22745249]
7. Bergethon K, et al. ROS1 rearrangements define a unique molecular class of lung cancers. J Clin Oncol. 2012; 30:863–870. [PubMed: 22215748]
8. Rikova K, et al. Global survey of phosphotyrosine signaling identifies oncogenic kinases in lung cancer. Cell. 2007; 131:1190–1203. [PubMed: 18083107]
9. Takeuchi K, et al. RET, ROS1 and ALK fusions in lung cancer. Nature medicine. 2012; 18:378–381.
10. Lundberg AS, et al. Immortalization and transformation of primary human airway epithelial cells by gene transfer. Oncogene. 2002; 21:4577–4586. [PubMed: 12085236]
11. Soda M, et al. Identification of the transforming EML4-ALK fusion gene in non-small-cell lung cancer. Nature. 2007; 448:561–566. [PubMed: 17625570]
12. Ju YS, et al. A transforming KIF5B and RET gene fusion in lung adenocarcinoma revealed from whole-genome and transcriptome sequencing. Genome research. 2012; 22:436–445. [PubMed: 22194472]
13. Kohno T, et al. KIF5B-RET fusions in lung adenocarcinoma. Nature medicine. 2012; 18:375–377.
14. Lipson D, et al. Identification of new ALK and RET gene fusions from colorectal and lung cancer biopsies. Nature medicine. 2012; 18:382–384.
15. Richardson C, Jasin M. Frequent chromosomal translocations induced by DNA double-strand breaks. Nature. 2000; 405:697–700. [PubMed: 10864328]
16. Lee HJ, Kweon J, Kim E, Kim S, Kim JS. Targeted chromosomal duplications and inversions in the human genome using zinc finger nucleases. Genome research. 2012; 22:539–548. [PubMed: 22183967]
17. Piganeau M, et al. Cancer translocations in human cells induced by zinc finger and TALE nucleases. Genome research. 2013; 23:1182–1193. [PubMed: 23568838]
18. Stephens PJ, et al. Massive genomic rearrangement acquired in a single catastrophic event during cancer development. Cell. 2011; 144:27–40. [PubMed: 21215367]
19. Baca SC, et al. Punctuated evolution of prostate cancer genomes. Cell. 2013; 153:666–677. [PubMed: 23622249]
20. Ran FA, et al. Double nicking by RNA-guided CRISPR Cas9 for enhanced genome editing specificity. Cell. 2013; 154:1380–1389. [PubMed: 23992846]
21. Mali P, et al. CAS9 transcriptional activators for target specificity screening and paired nickases for cooperative genome engineering. Nature biotechnology. 2013; 31:833–838.
22. Fu Y, Sander JD, Reyon D, Cascio VM, Joung JK. Improving CRISPR-Cas nuclease specificity using truncated guide RNAs. Nature biotechnology. 2014
23. Bassing CH, Swat W, Alt FW. The mechanism and regulation of chromosomal V(D)J recombination. Cell. 2002; 109(Suppl):S45–55. [PubMed: 11983152]

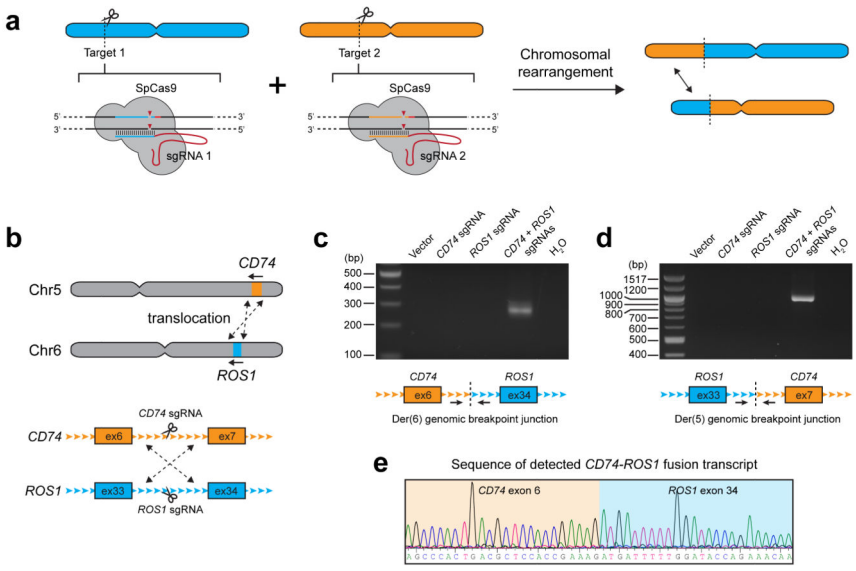

Figure 1. Cas9-induced DNA breaks promote interchromosomal translocations

a) Schematic depicting the overall strategy for generating chromosomal rearrangements. Cas9 from *S.pyogenes* (SpCas9) is co-expressed with two single-guide RNAs (sgRNA 1 and 2) which direct DNA cleavage at each targeted genomic site. b) The *CD74-ROS1* rearrangement results from a translocation between chromosomes 5 and 6. Shown are the intronic sites where Cas9 was targeted. c-d) PCR detection of the c) *CD74-ROS1* Der(6) and d) *ROS1-CD74* Der(5) genomic breakpoint junctions from HEK 293T cells in which Cas9 was expressed with no sgRNA (vector), *CD74* sgRNA alone, *ROS1* sgRNA alone, or both *CD74* and *ROS1* sgRNAs. e) Sequence chromatogram of the detected *CD74-ROS1* fusion transcript from cells in which Cas9 and both *CD74* and *ROS1* sgRNAs were expressed. Data shown are representative results from a total of three independent experiments.

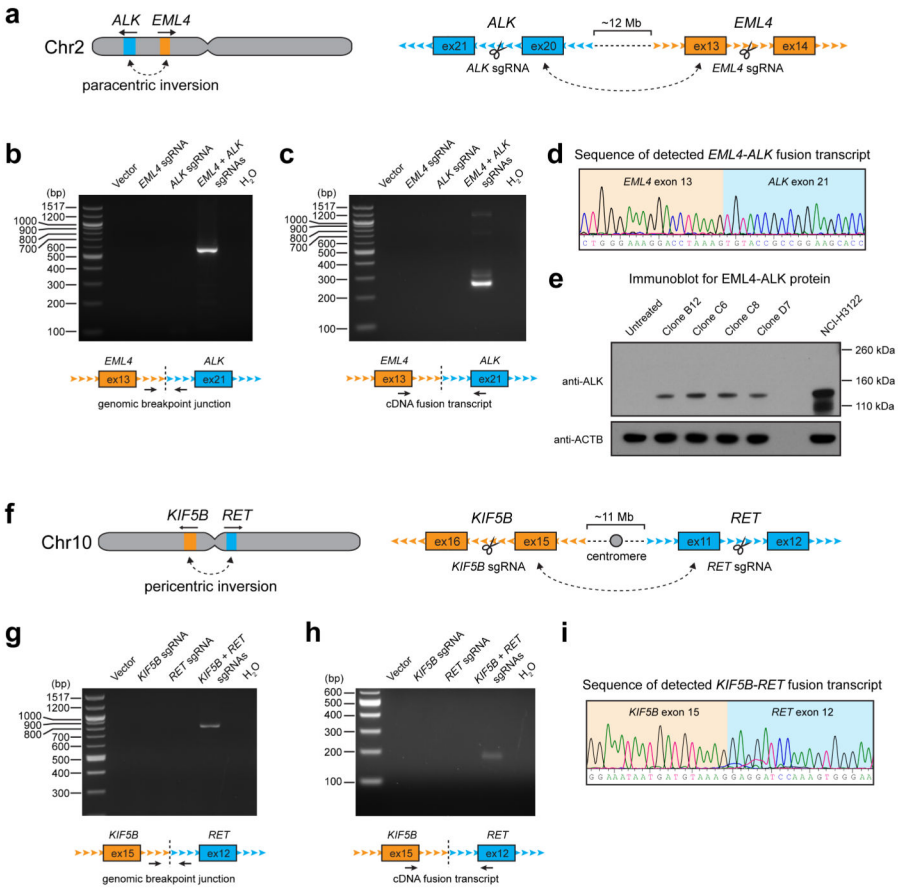

Figure 2. Cas9 can be targeted to generate paracentric and pericentric intrachromosomal inversions

a) The *EML4-ALK* rearrangement results from a paracentric inversion in chromosome 2. Shown are the intronic sites where Cas9 was targeted. b-c) PCR detection of the *EML4-ALK* b) genomic breakpoint junction and c) fusion transcript from HEK 293T cells in which Cas9 was expressed with no sgRNA (vector), *EML4* sgRNA alone, *ALK* sgRNA alone, or both *EML4* and *ALK* sgRNAs. d) Sequence chromatogram of the detected *EML4-ALK* fusion transcript from cells in which Cas9 and both *EML4* and *ALK* sgRNAs were expressed. e) Western blot for EML4-ALK protein expression in single clones of 293T cells which were untreated, or in which Cas9 and both *EML4* and *ALK* sgRNAs were expressed. NCIH3122 cells are shown as a positive control. f) The *KIF5B-RET* rearrangement results from a pericentric inversion in chromosome 10. Shown are the intronic sites where Cas9 was targeted. g-h) PCR detection of the *KIF5B-RET* g) genomic breakpoint junction and h) fusion transcript from HEK 293T cells in which Cas9 was expressed with no sgRNA (vector), *KIF5B* sgRNA alone, *RET* sgRNA alone, or both *KIF5B* and *RET* sgRNAs. i) Sequence chromatogram of the detected *KIF5B-RET* fusion transcript from cells in which Cas9 and both *KIF5B* and *RET* sgRNAs were expressed. Data shown are representative results from a total of three independent experiments.

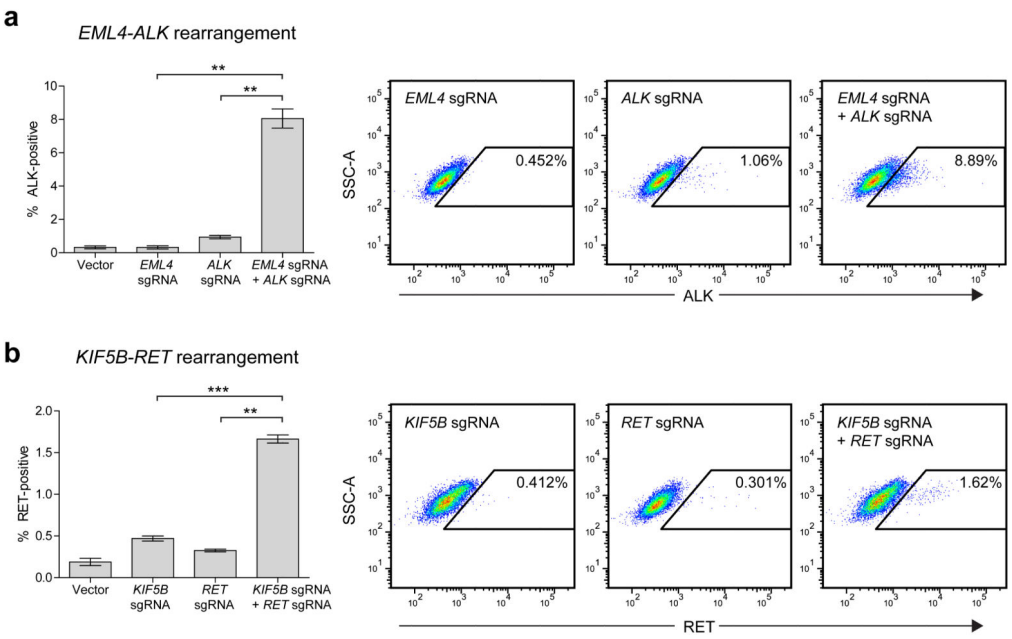

Figure 3. Estimated efficiencies of inducing *EML4-ALK* and *KIF5B-RET* rearrangements

HEK 293T cells transfected with Cas9 and no sgRNA (Vector), a single sgRNA, or a pair of sgRNAs were stained for either (a) ALK or (b) RET protein to estimate the percentage of cells with induced *EML4-ALK* or *KIF5B-RET* rearrangements, respectively. All values represent means of at least 3 independent experiments +/− s.e.m. Means were compared by paired t-test (**p<0.01, ***p<0.001). Also shown are dotplots from a representative experiment for each rearrangement.